

LE THÉORÈME DE KOMLÓS

Trois niveaux de difficulté/longueur :

- Piste bleue : partie 1.
- Piste rouge : parties 1 et 2.
- Piste noire : tout le devoir. La partie 3 est franchement subtile ! Si vous vous y frottez, merci de ne pas rédiger vos réponses à la partie 1.

Dans ce devoir, n est un entier naturel non nul. Soit M une matrice aléatoire de $\mathcal{M}_n(\mathbb{R})$ dont les coefficients sont des variables de Rademacher indépendantes. Le *théorème de Komlós* (1967), qui est l'objectif de ce devoir, énonce que :

$$P(M \notin \text{GL}_n(\mathbb{R})) \xrightarrow{n \rightarrow +\infty} 0.$$

Peu de temps après, Komlós a établi un résultat plus fin : $P(M \notin \text{GL}_n(\mathbb{R})) \underset{n \rightarrow +\infty}{=} O\left(\frac{1}{\sqrt{n}}\right)$. On sait depuis 2018 que $P(M \notin \text{GL}_n(\mathbb{R})) \underset{n \rightarrow +\infty}{=} \left(\frac{1}{2} + o(1)\right)^n$ et un chercheur prétend avoir montré en 2020 que $P(M \notin \text{GL}_n(\mathbb{R})) \underset{n \rightarrow +\infty}{\sim} \frac{2n^2}{2^n}$, mais je ne sais pas si sa preuve a été validée par les spécialistes du domaine.

La partie 1 est indépendante des autres. Les résultats de la partie 2 sont en revanche utilisés dans la partie 3.

1 MISE EN BOUCHE

- 1) Soit M une matrice aléatoire de $\mathcal{M}_2(\mathbb{C})$ dont les coefficients sont des variables aléatoires indépendantes uniformes sur \mathbb{U}_r — pour $r = 2$, les coefficients suivent donc la loi de Rademacher. Calculer $P(M \notin \text{GL}_2(\mathbb{C}))$.
- 2) a) Soit (Ω, P) un espace probabilisé fini. Montrer que pour tous événements A_1, \dots, A_n , en posant $\mathcal{P} = \mathcal{P}_2(\llbracket 1, n \rrbracket)$:

$$P\left(\bigcup_{i=1}^n A_i\right) \geq \sum_{i=1}^n P(A_i) - \sum_{\{i,j\} \in \mathcal{P}} P(A_i \cap A_j).$$

À présent, soit M une matrice aléatoire de $\mathcal{M}_n(\mathbb{R})$ dont les coefficients sont des variables de Rademacher indépendantes. On note C_1, \dots, C_n les colonnes de M .

- b) Calculer $P(C_i = C_j)$ et montrer que $P(C_i = C_j, C_k = C_l) = \frac{1}{4^n}$ pour tous $i, j, k, l \in \llbracket 1, n \rrbracket$ pour lesquels $i < j, k < l$ et $\{i, j\} \neq \{k, l\}$.
- c) En déduire que $P(M \notin \text{GL}_n(\mathbb{R})) \geq P\left(\bigcup_{1 \leq i < j \leq n} \{C_i = C_j\}\right) \underset{n \rightarrow +\infty}{\sim} \frac{n^2}{2^{n+1}}$.

Le résultat obtenu montre que la conjecture de l'introduction est essentiellement optimale.

2 THÉORÈMES DE SPERNER ET ERDÖS-LITTLEWOOD-OFFORD

Le *théorème d'Erdős-Littlewood-Offord* sera l'un des ingrédients majeurs de la preuve du théorème de Komlós.

- 3) Soit \mathcal{X} une *antichaine* de $\llbracket 1, n \rrbracket$, i.e. un ensemble de parties de $\llbracket 1, n \rrbracket$ dont aucune ne contient aucune autre. Soit par ailleurs σ une permutation aléatoire de $\llbracket 1, n \rrbracket$. Pour tout $X \in \mathcal{X}$, on note B_X l'événement « $\sigma|_{\llbracket 1, |X| \rrbracket}$ est bijective de $\llbracket 1, |X| \rrbracket$ sur X ».
 - a) Calculer $P(B_X)$ pour tout $X \in \mathcal{X}$.
 - b) En déduire l'inégalité *LYM* ou *inégalité de Lubell-Yamamoto-Meshalkin* : $\sum_{X \in \mathcal{X}} \binom{n}{|X|}^{-1} \leq 1$.
 - c) En déduire le *théorème de Sperner* : $|\mathcal{X}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.
- 4) Montrer que la suite $\left(\frac{\sqrt{n+1}}{2^{2n}} \binom{2n}{n}\right)_{n \in \mathbb{N}}$ décroît, puis en déduire que pour tout $n \in \mathbb{N}^*$: $\binom{n}{\lfloor \frac{n}{2} \rfloor} \leq 2^n \sqrt{\frac{2}{n}}$.

- 5) Soient $\varepsilon_1, \dots, \varepsilon_n$ des variables de Rademacher indépendantes, $a_1, \dots, a_n > 0$ et $x \in \mathbb{R}$.
- a) Montrer que l'ensemble $\left\{ X \in \mathcal{P}(\llbracket 1, n \rrbracket) \mid \sum_{i \in X} a_i - \sum_{j \notin X} a_j = x \right\}$ est une antichaîne de $\llbracket 1, n \rrbracket$.
- b) En déduire que : $P(\varepsilon_1 a_1 + \dots + \varepsilon_n a_n = x) \leq \sqrt{\frac{2}{n}}$ (théorème d'Erdős-Littlewood-Offord).
- c) Montrer que l'inégalité **b)** reste vraie sous l'hypothèse que $a_1, \dots, a_n \in \mathbb{R}^*$.

3 LE THÉORÈME DE KOMLÓS

- 6) Soit V un sous-espace vectoriel de dimension d de \mathbb{R}^n . On se donne une matrice $A \in \mathcal{M}_{n,d}(\mathbb{R})$ dont les colonnes forment une base de V . Pourquoi A possède-t-elle une matrice extraite inversible de taille d ? En déduire que :

$$|V \cap \{-1, 1\}^n| \leq 2^d.$$

- 7) Soient $\lambda \in \left[\frac{1}{2}, 1 \right[$ et X une variable aléatoire de loi $\mathcal{B}\left(n, \frac{1}{2}\right)$.
- a) Montrer que pour tout $t \geq 0$: $e^{n\lambda t} P(X \geq \lambda n) \leq \left(\frac{e^t + 1}{2}\right)^n$.
- b) Montrer que la fonction $t \mapsto e^{-\lambda t} (e^t + 1)$ admet le réel $\mu = (\lambda^\lambda (1-\lambda)^{1-\lambda})^{-1}$ pour minimum sur \mathbb{R}_+ .
- c) En déduire que $\sum_{0 \leq k \leq (1-\lambda)n} \binom{n}{k} \leq \mu^n$, puis que $\sum_{0 \leq k \leq \frac{n}{5}} \binom{n}{k} 2^k \leq \eta^n$ avec $\eta = 5 \cdot 2^{-\frac{7}{5}}$. On admet que $\eta < 2$, ce sera très important à la fin.

Pour une raison technique liée à la question **10)c)**, les espaces vectoriels manipulés ci-dessous sont des \mathbb{Q} -espaces vectoriels.

Soit M une matrice aléatoire de $\mathcal{M}_n(\mathbb{Q})$ dont les coefficients sont des variables de Rademacher indépendantes. On note C_1, \dots, C_n les colonnes de M et on pose $V_j = \text{Vect}(C_1, \dots, C_j)$ pour tout $j \in \llbracket 1, n \rrbracket$. Pour tout sous-espace vectoriel W de \mathbb{Q}^n , on pose enfin $W^\perp = \{X \in \mathbb{Q}^n \mid \forall Y \in W, Y^\top X = 0\}$.

- 8) Montrer que $P(C_{j+1} \in V_j) \leq \frac{1}{2^{n-j}}$ pour tout $j \in \llbracket 1, n-1 \rrbracket$, puis que pour tout $k \in \llbracket 1, n-1 \rrbracket$:

$$P(M \notin \text{GL}_n(\mathbb{R})) \leq \frac{1}{2^k} + \sum_{j=n-k}^{n-1} P(C_{j+1} \in V_j).$$

On note E (resp. E_3) l'ensemble des vecteurs non nuls de \mathbb{Q}^n (resp. $\{0, 1, 2\}^n$) dont au plus $\frac{n}{5}$ coefficients sont non nuls.

- 9) a) Montrer que $|E_3| \leq \eta^n$, où le réel η est celui de la question **7)c)**.
- b) Montrer que pour tout $j \in \llbracket 1, n \rrbracket$: $\{V_j^\perp \cap E \neq \emptyset\} \subset \bigcup_{Y \in E_3} \bigcap_{1 \leq i \leq j} \{C_i^\top Y \equiv 0 [3]\}$.
- c) Soient $Y \in \{0, 1, 2\}^n$ non nul et X un vecteur aléatoire de \mathbb{Q}^n dont les coefficients sont des variables de Rademacher indépendantes. Pour tout $i \in \llbracket 1, n \rrbracket$, on note $X_{(i)}$ le vecteur obtenu à partir de X par simple multiplication par -1 du $i^{\text{ème}}$ coefficient. Montrer que les événements $\{X^\top Y \equiv 0 [3]\}$ et $\{X_{(i)}^\top Y \equiv 0 [3]\}$ sont incompatibles pour un certain $i \in \llbracket 1, n \rrbracket$ et en déduire que $P(X^\top Y \equiv 0 [3]) \leq \frac{1}{2}$.
- d) En déduire que pour tout $j \in \llbracket 1, n \rrbracket$: $P(V_j^\perp \cap E \neq \emptyset) \leq \frac{\eta^n}{2^j}$.

- 10) Soit $j \in \llbracket 1, n-1 \rrbracket$. On note \mathcal{W}_j l'ensemble des sous-espaces vectoriels de \mathbb{Q}^n engendrés par j vecteurs de $\{-1, 1\}^n$ pour lesquels $W^\perp \cap E = \emptyset$.

- a) Montrer que $W^\perp \neq \{0\}$ pour tout $W \in \mathcal{W}_j$.
- b) Les coefficients $m_{1,j+1}, \dots, m_{n,j+1}$ sont des variables de Rademacher indépendantes au sens de la probabilité P . Pourquoi cela reste-t-il vrai pour tout $W \in \mathcal{W}_j$ au sens de la probabilité conditionnelle $P_{\{V_j=W\}}$?
- c) En déduire que $P_{\{V_j=W\}}(C_{j+1} \in V_j) \leq \sqrt{\frac{10}{n}}$ pour tout $W \in \mathcal{W}_j$, puis que $P(C_{j+1} \in V_j) \leq \sqrt{\frac{10}{n}} + \frac{\eta^n}{2^j}$.

- 11) En exploitant le résultat de la question **8)** pour $k = \left\lfloor \frac{\ln n}{\ln 2} \right\rfloor$, montrer finalement que $P(M \notin \text{GL}_n(\mathbb{R})) \underset{n \rightarrow +\infty}{=} O\left(\frac{\ln n}{\sqrt{n}}\right)$.